

Package ‘gsalib’

February 20, 2015

Type Package

Title Utility Functions For GATK

Version 2.1

Date 2014-12-09

Author Kiran Garimella

Maintainer Geraldine Van der Auwera <vdauwera@broadinstitute.org>

Description

This package contains utility functions used by the Genome Analysis Toolkit (GATK) to load tables and plot data. The GATK is a toolkit for variant discovery in high-throughput sequencing data.

License MIT + file LICENSE

LazyLoad yes

NeedsCompilation no

Repository CRAN

Date/Publication 2014-12-10 08:09:01

R topics documented:

gsalib-package	2
gatkreport_test_v1.table	3
gsa.read.gatkreport	4
gsa.read.gatkreportv0	5
gsa.read.gatkreportv1	5

Index	7
--------------	----------

gsalib-package

Utility functions for GATK

Description

Utility functions for analysis of genome sequence data with the GATK

Details

Package: gsalib
Type: Package
Version: 2.1
Date: 2014-12-09
License: MIT
LazyLoad: yes

This package is primarily meant to be used programmatically by GATK tools. However the `gsa.read.gatkreport()` function can be used to easily read in data from a GATKReport. A GATKReport is a multi-table document generated by GATK tools.

Author(s)

Kiran Garimella

Maintainer: Geraldine Van der Auwera <vdauwera@broadinstitute.org>

References

<http://www.broadinstitute.org/gatk/guide/article?id=1244>

Examples

```
test_file = system.file("inst", "extdata", "test_gatkreport.table", package = "gsalib");  
report = gsa.read.gatkreport(test_file);
```

gatkreport_test_v1.table

Test table for version 1.x GATKReport

Description

This is a new-style GATKReport.

Usage

```
gatkreport_test_v1.table
```

Format

Text document containing multiple tables

Source

GSA test data, Broad Institute

References

<http://www.broadinstitute.org/gatk/guide/article?id=1244>

gsa.read.gatkreport *Function to read in a GATKReport*

Description

This function reads in data from a GATKReport. A GATKReport is a document containing multiple tables produced by the GATK. Each table is loaded as a separate data.frame object in a list.

Usage

```
gsa.read.gatkreport(filename)
```

Arguments

filename The path to the GATKReport file.

Details

The GATKReport format replaces the multi-file output format used by many GATK tools and provides a single, consolidated file format. This format accommodates multiple tables and is still R-loadable through this function.

Value

Returns a LIST object, where each key is the TableName and the value is the data.frame object with the contents of the table. If multiple tables with the same name exist, each one after the first will be given names of TableName.v1, TableName.v2, ..., TableName.vN.

Note

This function accepts different versions of the GATKReport format by making internal calls to gsa.read.gatkreportv0() or gsa.read.gatkreportv1() as appropriate.

Author(s)

Kiran Garimella

References

<http://www.broadinstitute.org/gatk/guide/article?id=1244>

Examples

```
test_file = system.file("inst", "extdata", "test_gatkreport.table", package = "gsalib");
report = gsa.read.gatkreport(test_file);
```

`gsa.read.gatkreportv0` *Function to read in an old-style GATKReport*

Description

This function reads in data from a version 0.x GATKReport. It should not be called directly; instead, use `gsa.read.gatkreport()`

Usage

```
gsa.read.gatkreportv0(lines)
```

Arguments

`lines` The lines read in from the input file.

Value

Returns a LIST object, where each key is the `TableName` and the value is the `data.frame` object with the contents of the table. If multiple tables with the same name exist, each one after the first will be given names of `TableName.v1`, `TableName.v2`, ..., `TableName.vN`.

Author(s)

Kiran Garimella

References

<http://www.broadinstitute.org/gatk/guide/article?id=1244>

`gsa.read.gatkreportv1` *Function to read in a new-style GATKReport*

Description

This function reads in data from a version 1.x GATKReport. It should not be called directly; instead, use `gsa.read.gatkreport()`

Usage

```
gsa.read.gatkreportv1(lines)
```

Arguments

`lines` The lines read in from the input file.

Value

Returns a LIST object, where each key is the TableName and the value is the data.frame object with the contents of the table. If multiple tables with the same name exist, each one after the first will be given names of TableName.v1, TableName.v2, ..., TableName.vN.

Author(s)

Kiran Garimella

References

<http://www.broadinstitute.org/gatk/guide/article?id=1244>

Index

*Topic **manip**

- gsa.read.gatkreport, 4
- gsa.read.gatkreportv0, 5
- gsa.read.gatkreportv1, 5

*Topic **package**

- gsalib-package, 2

gatkreport_test_v1.table, 3

gsa.read.gatkreport, 4

gsa.read.gatkreportv0, 5

gsa.read.gatkreportv1, 5

gsalib (gsalib-package), 2

gsalib-package, 2